

McNamara Declaration

Exhibit 24

1 IN THE UNITED STATES DISTRICT COURT FOR THE SOUTHERN
2 DISTRICT OF NEW YORK

3
4 HACHETTE BOOK GROUP, INC.,) Case No.
HARPERCOLLINS PUBLISHERS LLC,) 1:20-CV-04160-JGK
5 JOHN WILEY & SONS, INC., and)
PENGUIN RANDOM HOUSE, LLC,)
6)
Plaintiffs,)
7 v.)
INTERNET ARCHIVE and DOES 1)
8 through 5, inclusive,)
)
9 Defendants.)
-----)

10
11
12
13
14 REMOTE PROCEEDINGS OF THE
15 VIDEOTAPED DEPOSITION OF BRENTON CHENG,
16 INDIVIDUALLY AND AS A 30(B)(6) WITNESS
17 FOR INTERNET ARCHIVE
18 FRIDAY, DECEMBER 3, 2021
19
20
21
22

23 REPORTED BY NANCY J. MARTIN
24 CSR. NO. 9504, RPR, RMR
25 PAGES 1 - 272

1 corrected color, you know, the cleaned up images that
2 come from the book are then used to -- during the
3 derive process is where the OCR is run, optical
4 character recognition. It's where the underlying PDF
5 or ePUB gets created. A bunch of metadata gets
6 adjusted and put into its right place. It's
7 essentially the core processing pipeline for the book.

8 Q. So the output of the derive process results
9 in the initial scan of the book after it's been --
10 sorry. Strike that.

11 The output of the derive process results in
12 the first digitized scan of the book that the Internet
13 Archive has made being turned into copies in various
14 other forms; is that correct?

15 A. That's right, among other things. Metadata,
16 various other changes, supplemental data, that kind of
17 thing.

18 Q. And so the scan that the Internet Archive has
19 made of a physical book is in a JPEG format initially,
20 and then the Internet Archive, you know, after it's
21 done what it has done to crop the book and adjust
22 metadata, you know, and what not, it then creates
23 copies of the book in various formats, including ePUB,
24 PDF, DJVU, several other forms; correct?

25 A. That's right. If you're looking for one

1 word, the terminology of that clean-up process, we
2 call it republishing.

3 Q. So the Internet Archive views itself as
4 republishing the physical book but as a digital
5 version; correct?

6 A. Yeah.

7 MS. LEE: Sorry, Brenton. Objection.
8 Mischaracterizes testimony.

9 THE WITNESS: Yes. Okay. Maybe that wasn't
10 helpful. But it's an internal term. It's not
11 connected with -- not intended to connect with any
12 other meaning around publishing. It's simply an
13 internal term that we use to summarize the book image
14 correction process. Maybe I should just stick with
15 that.

16 BY MR. ZEBRAK:

17 Q. So the output of the derive process is that
18 the Internet Archive is putting into its -- onto its
19 website the digital version that the Internet Archive
20 has made of the physical book; is that correct?

21 A. Right. To the degree that one has access to
22 it. So on the website and our servers, and then
23 access depends on what the user has access to.

24 Q. So how many different -- so the initial JPEG
25 scan of the book that the Internet Archive makes, how

C E R T I F I C A T E

I do hereby certify that the aforesaid testimony was taken before me, pursuant to notice, at the time and place indicated; that said deponent was by me duly sworn to tell the truth, the whole truth, and nothing but the truth; that the testimony of said deponent was correctly recorded in machine shorthand by me and thereafter transcribed under my supervision with computer-aided transcription; that the deposition is a true and correct record of the testimony given by the witness; and that I am neither of counsel nor kin to any party in said action, nor interested in the outcome thereof.



Nancy J. Martin, RMR, CSR

Dated: December 6, 2021

(The foregoing certification of this transcript does not apply to any reproduction of the same by any means, unless under the direct control and/or supervision of the certifying shorthand reporter.)